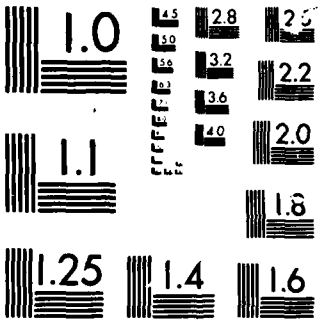
AD-A166 966 DEVELOPMENT OF A COMPUTER-BASED WRITING ASSISTANT: REVISE(U) NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER SAN DIEGO CA F R CHANG MAR 86 NPRDC-TR-86-10 1/1

UNCLASSIFIED F/G 9/2 NL



MICROCOPY CHART

NPRDC TR 86-10

MARCH 1986

DEVELOPMENT OF A COMPUTER-BASED WRITING ASSISTANT: REVISE

NAVY PERSONNEL RESEARCH
AND
DEVELOPMENT CENTER
San Diego, California 92152

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY PRACTICABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

NPRDC TR 86-10

March 1986

DEVELOPMENT OF A COMPUTER-BASED WRITING ASSISTANT: REVISE

Frederick R. Chang

Reviewed by
E. G. Aiken

Approved by
J. C. McLachlan

Released by
Howard S. Eldredge
Captain, U.S. Navy
Commanding Officer

Approved for public release; distribution unlimited.

Navy Personnel Research and Development Center
San Diego, California 92152-6800

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

ADA 166966

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NPRDC TR 86-10	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Navy Personnel Research and Development Center 6b. OFFICE SYMBOL (If applicable) Code 51	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) San Diego, CA 92152-6800	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Chief of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 63720N; PROJECT NO: R-1772; TASK NO: ; WORK UNIT ACCESSION NO: 003

11. TITLE (Include Security Classification)

DEVELOPMENT OF A COMPUTER-BASED WRITING ASSISTANT: REVISE

12. PERSONAL AUTHOR(S)

Chang, F. R.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 83 Oct TO 84 Sep	1986 March	24

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	
05	09		Clear writing, text design, written communication, computers and writing, writing improvement

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This paper describes the development of an efficient, easy-to-use computer-based writing assistant, called REVISE, that implements some of the plain English guidelines specified in SECNAVINST 5216.5C. It discusses the human-computer interface problems of some existing computer-based writing systems and how REVISE addresses these problems. A preliminary evaluation of REVISE is described, as are directions for future research and development on computer-based writing aids.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL F. R. Chang	22b. TELEPHONE (Include Area Code) (619) 225-6434 22c. OFFICE SYMBOL Code 51

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

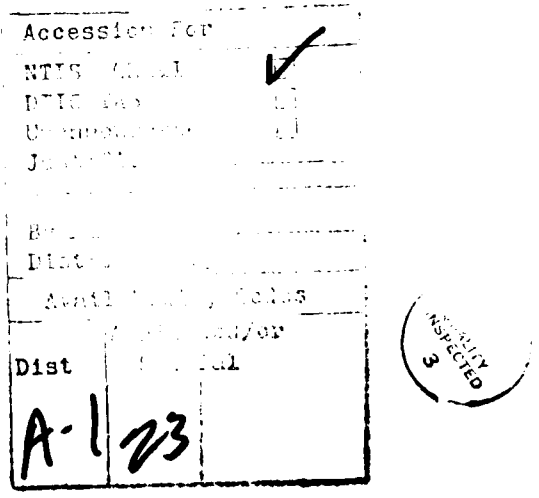
FOREWORD

This research was conducted within the in-house independent laboratory research program. The objective of this work unit was to begin development of useful and efficient computer-based writing tools.

This research should be of interest to developers of computer-based writing aids and researchers concerned with effective writing.

H. S. ELDREDGE
Captain, U.S. Navy
Commanding Officer

J. W. TWEEDDALE
Technical Director



SUMMARY

Problem and Background

In recent years, interest in improving the communication efficiency of written documents has been increasing. Although numerous guidelines suggest ways to improve government and Department of Defense writing, the problems persist. Possibly, the guidelines are not being read and followed, the procedures are too difficult and time-consuming to implement, or the suggestions themselves are not adequate to produce the desired results.

Evidence now available suggests that people have great difficulty detecting problem areas in a text even when they know what they are supposed to detect. Apparently, most people cannot "turn off" normal comprehension processes to detect the problems. A computerized writing system could detect these problem areas much better than humans. A few computerized writing assistants are available; however, users may find their output difficult to use and not particularly helpful and interaction with the programs somewhat inefficient.

Existing guidelines for improving the writing in documents include the latest Navy Correspondence Manual (SECNAVINST 5216.5C), which is of particular interest to the Navy. The first chapter, which advocates the use of plain English, details fairly concrete steps for improving communication efficiency such as using short, simple, and common words and phrases and active (versus passive) sentence constructions.

Objective

The objective of this effort was to develop an efficient, easy-to-use computer-based writing assistant (REVISE) to help writers follow some of the plain English writing guidelines specified in SECNAVINST 5216.5C.

Approach

Development of REVISE followed the design guidelines that a computer-based writing assistant system should: (1) be easy to use, (2) identify specifically what is wrong with the writing in the document, (3) provide specific recommendations for improving the writing, (4) have a quick and relatively inexpensive user interface, and (5) run on mainframe computers as well as on smaller microcomputers.

Data Base. The SECNAVINST 5216.5C lists commonly used complex words and phrases with simpler alternatives. This list with modifications and enhancement formed the basis for REVISE's data base.

Operation. REVISE runs interactively and analyzes each sentence in the document separately. When it detects a word or phrase that is in the data base, it displays the entire sentence with the word or phrase highlighted and lists suggested alternatives to replace the highlighted items. The system brings up a menu of options for changing the document quickly and easily. Once the revision is completed, the system creates a new file containing all changes. The original file is left intact.

System Evaluation. An untrained user used REVISE to analyze and revise 36 documents. Several surface structure characteristics were measured and compared.

Results

In general, over 90 percent of the words and phrases flagged by the system were appropriate to the sentence context. When various text characteristics of the original and revised documents were compared, the revised documents showed significant improvements on measures thought to influence communication efficiency. For example, the revised documents generally used simpler and more commonly used words, fewer passive constructions, and fewer nominalizations. Sentence construction measurements showed no significant changes.

Conclusions

REVISE improved the communication efficiency of 36 Navy documents. The system described here represents an efficient computerized writing assistant that can help a writer follow plain English guidelines.

Recommendation

Design of future computer-based writing assistants should be based on an explicit model of comprehension derived from empirical findings from the psycholinguistic literature combined with the efficient user interface of REVISE.

CONTENTS

LIST OF TABLES

INTRODUCTION

Problem

Many people in important contexts support the increasing interest in improving the communication efficiency of written documents. For example, in a letter to Navy Undersecretary Goodrich, Vice Admiral Mustin stated that "The fleet is outraged over poor technical manual quality" (Simply Stated, October 1984).

A number of state legislatures have imposed legal requirements for the use of "plain English" in public documents. The courts are also involved. Recently, for example, a federal judge in Brooklyn, New York ordered the government to remove the "gobbledygook" from Medicare benefit form letters. An elderly, retired man could not understand why identical claims submitted for his wife's treatments resulted in different payments. The long and confusing letters he received from Medicare angered him so much that he took his case to small claims court. In a class-action suit, the judge ruled that the language used in the letters could not be understood by their recipients and, in fact, "did not qualify as English" (Simply Stated, July 1984).

Background

Government writing has been the target of much criticism through the years (hence, the term "governmentese"). Despite the numerous guidelines suggesting ways to improve government and Department of Defense (DoD) writing (Department of the Air Force, 1973; Department of the Interior, 1978; Department of the Treasury, 1969, 1975; Secretary of the Navy, 1983), the problems persist. Apparently, the guidelines are not being read and followed, the procedures are too difficult and time-consuming to implement, or the suggestions are not adequate to produce the desired results.

Wright (1985) suggests that people have great difficulty detecting problems in a text. She gave two groups of subjects who had no formal writer training some documents to revise. One group received a few guidelines to follow (e.g., avoid passive verbs and complex words); and the other, did not. Although the guideline group detected more problems than the no-guideline group, they detected less than 40 percent of the problems addressed in the guidelines. This result suggests that people do not detect problems in writing well, probably at least partly because reading for most people is highly automatized (LaBerge & Samuels, 1974; Posner & Snyder, 1974). Thus, people find it hard to "turn off" the normal comprehension processes in order to detect problems in the text.

Kieras (1985) considered the error detection problem to be an important point because current computer technology might aid this human capability. A computerized writing system could detect the problems much better than humans could. Thus, a well-designed system might benefit writers of informational text greatly.

Presently two major computer-based writing aids[1] are available: (1) the Writer's Workbench™ (WWB) developed at Bell Laboratories (Cherry, 1982; Macdonald, 1983; Macdonald, Frase, Gingrich, & Keenan, 1982) and (2) the Computer Readability Editing Systems (CRES) developed by the Navy (Kincaid, Aagard, & O'Hara, 1980; Kincaid, Aagard, O'Hara, & Cottrell, 1981; Kincaid, Cottrell, Aagard, & Risley, 1981). We will discuss these two systems briefly, along with some issues connected with the human-computer interface.

---

[1] Identification of equipment and software is for documentation only and does not imply endorsement.

Writer's Workbench™ (WWB)

The WWB is a set of 28 separate computer programs designed to help a writer improve a text. The system runs on the UNIX™ operating system, and each program is called by executing a separate UNIX command. The heart of the system is a program that determines parts of speech in a text using a rule-based approach (Cherry, 1978). Many of the WWB programs are designed around this parts-of-speech program. One of the programs, for example, provides the writer or user with a "scorecard" of the frequency of use of a variety of different word types (e.g., nouns, adjectives, adverbs). The WWB system provides programs for the user to check spelling, punctuation, diction, sexist language, readability level, and abstractness.

Overall, the WWB system's impressive program diversity makes it quite useful for some purposes. However, the system has some important problems, stemming mainly from the system's interface with the user. Perhaps the more fundamental problem is the amount of effort that users must expend to use the information provided by WWB. For example, to check the spelling in a document, the user calls the spell checker on a document and the program lists the misspelled items on the user's screen. (The output can be sent to a file and stored, if the user knows the additional UNIX commands to do so.) Next, the user writes down the misspellings (or requests a printed output), calls the text editor on the document, locates the misspelled words in the document, corrects the misspelled words using a dictionary as needed (to query the system dictionary for spelling the user must run another program and record the output), and then saves the file. Clearly, this is time consuming and not very efficient.

Another problem with WWB concerns the specificity of the problems found by the programs and the suggested modifications. For example, telling the user that the readability score of the document is too high and to decrease it is not too helpful. The document "scorecard," mentioned earlier, is very difficult to interpret and figuring out how to modify the document is even more so.

Finally, the WWB system is difficult to use. Since WWB was designed to run under the UNIX operating system, it adopted many UNIX conventions. Thus, many of the programs have several "options" that cause the program to behave differently. Adoption of UNIX conventions provides WWB power, but at the expense of ease of use (Norman, 1981). The WWB user's manual is over 300 pages long and does not include any discussion of UNIX, per se. A working knowledge of UNIX greatly facilitates efficient use of WWB.

Computer Readability Editing System (CRES)

The CRES system was designed specifically to improve the comprehensibility of Navy technical manuals and instructional materials. CRES's three primary tasks are to: (1) flag uncommon, complex, or misspelled words and long and passive sentences in a document, (2) suggest alternatives for some of the flagged words, and (3) compute the document's readability score. CRES contains five published military and nonmilitary common word lists, one special list of Navy technical words, and a list of potential alternatives for some of the flagged words.

Originally, CRES was designed to run on a Wang word processor with a 64K main memory and was written in a dialect of BASIC. CRES is smaller than WWB and, therefore, considerably easier to use. However, as with WWB, users encounter some problems with implementing the suggestions CRES provides. Processing a document through CRES produces an output file that contains an annotated version of the actual

document. The comments include brackets around flagged words, alternatives after some of the bracketed words, identification of passive constructions (by surface structure rules) and sentences exceeding 22 words in length, and some statistics about the document (e.g., readability score). The output is useful, but, as with WWB output, the user must go over the output in detail, call up the word processor on the document, search for each flagged item in the document, make the substitution, and then save the file.

Other Computer-based Writing Aids

A few other systems are under development, notably the EPISTLE system from IBM (Heidorn, Jensen, Miller, Byrd, & Chodorow, 1980) and a promising system by Kieras (1985) at the University of Michigan based on artificial intelligence programming techniques. Neither system is generally available yet.

Writing Guidelines

A variety of guidelines to improve writing also exist. Of particular interest to the Navy is the latest Navy Correspondence Manual, (SECNAVINST 5216.5C) (Secretary of the Navy, 1983). Its first chapter reiterates the prevalence of poor writing, advocates using plain English, and details fairly concrete steps on how to improve writing. The steps include using simple, common words and phrases and constructing active (versus passive) sentences.

Objective

The primary objective of this effort was to develop an efficient, easy-to-use computer-based writing assistant (REVISE) to help writers follow the plain English writing guidelines specified in SECNAVINST 5216.5C. A secondary goal was a preliminary evaluation of whether or not changes made by the system improve writing quality.

APPROACH

System Design Issues

Design of the computer-based writing system--REVISE--was based on the following simple but powerful design principles: The system should (1) be easy to use, (2) identify the document's specific problems, (3) provide specific recommendations for improving writing in the document, (4) provide an efficient interface, and (5) run on mainframe and microcomputers. Each guideline is discussed in greater depth below.

1. Be easy to use. Perhaps the most important design principle was user-friendliness. REVISE was designed for users who have very little time to learn a system. In a recent review of the UNIX operating system, Norman (1981) pointed out the importance of making a system comprehensible to the user. No matter how technically sound the computational algorithms are, if the system is difficult to use, people may find it too tedious to use the system at all.

2. Identify the document's specific problems. The system should provide specific information about the writing in the document. For example, merely stating that the readability level of the document is too high is not specific enough. The output of one of the main WWB programs is a set of statistics about the document including how many nouns, verbs, and adjectives are in the document; four different readability scores; and

information about sentence and word length. The WWB documentation gives some general guidelines for interpreting output, but, in general, these are not specific enough to be helpful to the user.

3. Provide specific recommendations for improving the writing in the document. A related notion concerns giving the user specific information about how to improve the document. Giving specific information about problems with the document is not enough; specific suggestions for correcting the problem are needed. Research has shown that writers are not always able to correct a particular problem even if they know what the problem is (Hayes, 1984).

4. Provide an efficient interface. The system should also provide an efficient interface. Writers might use such a system if it saves time. Thus, a system that provides output about what is wrong with a document but makes the user re-edit the document in order to incorporate the changes is inefficient with the user's time.

5. Run on mainframe and microcomputers. Many writers have access to large mainframe systems that have sophisticated document preparation features. Others, though, have access only to microcomputer systems. Thus, to be useful to as many writers as possible, the system should be transportable to microcomputers as well as mainframe environments.

REVISE System Features

1. Data base. The major function of REVISE is to help writers use simple words and phrases with the aid of a data base. SECNAVINST 5216.5C lists awkward, uncommon, or complex words and phrases commonly used in Navy correspondence along with suggested alternatives. This list formed the basis for REVISE's data base with the items modified and enhanced for use in the system. Following SECNAVINST 5216.5C, the system checks for passive sentences (in a minimal way) by flagging constructions such as "it is" that tend to delay meaning, hide responsibility, and encourage passive verbs.

The data base contains 446 words and phrases that are flagged in a document and 723 simpler alternatives for these 446 items. The data base includes the past tense of verbs and plural endings. One issue of interest concerns the relative frequency of use in the English language of the complex items compared with the suggested, simpler alternatives. The comparison provides an objective way to determine the extent to which REVISE suggests simpler more common language. This frequency can be estimated through the use of one of several published word frequency lists. The list by Kucera and Francis (1967) revealed that the items in the data base (only single word entries were selected) had an average Kucera-Francis rating of 25.5 (this means that the words appeared in a sample of slightly over a million words, an average of 25.5 times). The simpler suggestions had a mean Kucera-Francis rating of 857.2. The difference between the means was statistically reliable. (In the analysis, all words not appearing on the list were given a value of 1.)

2. Operation. The writer can run the draft of any document entered in the computer--mainframe or microcomputer--through the REVISE program. The program runs interactively and analyzes each sentence sequentially. When the system detects a word or phrase contained in the data base, it displays the entire sentence, highlights the detected word or phrase, and displays alternative words or phrases that might replace the detected items. The system lists a menu of options that enable the user to: (a) change the detected word or phrase, (b) rewrite the entire sentence, (c) leave the sentence

unchanged, (d) display again the current sentence showing the problem and the change, (e) read several sentences that precede the current sentence, or (f) stop the program for now, saving all changes in a temporary file. The menu provides a quick, easy way to make changes to the document. Once the revision is completed, the system creates a file on disk that contains all changes. The original file is left intact. Because REVISE supports the standard UNIX document format programs, the revised document will be identical in format to the original and can be printed without further modifications. The system also gives statistics on the number of sentences analyzed, the number of words or phrases flagged, and the number of changes made.

3. The "C" language. The system was developed on the Navy Personnel Research and Development Center (NAVPERSRANDCEN) DEC 11/780 VAX mainframe running the UNIX (4.2 BSD) operating system. The program was written in the "C" programming language because it produces fast, compact, executable code on the VAX and is transportable to other computing environments. A version of the program also runs on the IBM PC microcomputer. The smallest version of the system will run in 64K of memory with one floppy disk drive.

System Evaluation Issues

Does REVISE improve writing quality? That is, do the plain English prescriptions implemented in REVISE actually influence characteristics of documents in ways that improve writing and communication efficiency? As the major objective of this development effort was to automate portions of SECNAVINST 5216.5C, a full-scale evaluation of whether REVISE actually improved comprehensibility (as assessed, say, through psychological experimentation), is beyond the scope of this report. However, we can assess the effectiveness of REVISE at certain levels.

REVISE can influence a document in a variety of ways. REVISE should decrease the length of the overall document, average sentence, or average word. Changes of this sort reflect the prescription to be brief. Since standard readability scores are a function of average sentence and word length, REVISE should increase the measured readability of a document and thus lower the reading grade level associated with it, compared to the original.

At a more specific level, we can examine how REVISE influences the frequencey with which particular word types are used. Since REVISE checks for expletive sentence beginnings ("it" and "there" with some form of the verb "to be" in constructions where the subject follows the verb, e.g., "There were too many people at the meeting"), they should be decreased. By checking for expletive beginnings, REVISE also checks for "to be" verbs; thus, the proportion of these verbs should go down. Similarly, since some expletive sentence beginnings are also associated with passive verbs (e.g., "It is hoped that the meeting will continue"), the proportion of passive verbs should decrease. A variety of other specific word types can be examined. Additionally, we could look at factors related to sentence complexity (such as the proportion of simple, compound, and complex sentences); however, since REVISE makes its suggestions primarily at the word and phrase level, we should not expect large changes at the sentence level.

An important question is whether or not REVISE detects appropriate document problems; that is, in the context of a particular sentence, is the flagged item appropriate? For example, many individual English words can be used as nouns or verbs. The context determines the appropriate word class. Because the items in the data base assume a particular word class, in some of the cases, the system may flag an inappropriate item.

For example, REVISE considers the verb "request" awkward compared to "ask"; however, "request" can also be used as a noun.

A simple methodology to assess some of these factors is to use REVISE to revise some actual Navy documents and compare the revisions with the originals. The procedure used is described below.

1. Documents. For this study, 36 Navy letters and memos were collected from the master file of documents from the NAVPERSRANDCEN mail room. The documents originated from a variety of naval commands including NAVPERSRANDCEN, Naval Material Command, Naval Education and Training Command, Office of the Chief of Naval Operations, and others.

2. Program user. An undergraduate college student, who did not know the purpose of the study and did not assist in the development of the program, used the REVISE system to analyze and modify the documents.

3. Procedure. The user was told that we had developed a program to assist writers in revising documents and to use the program to analyze the 36 documents. He was told to categorize each flagged item as (a) a single word, (b) a phrase, or (c) not appropriate to the sentence. The user was told to change the sentence based on the program's suggestions if he thought one of the suggestions improved the document.

4. Dependent variables. Several surface structure characteristics of the original and revised documents were measured: various document length measures, readability scores, specific word type measures, and measures of sentence complexity.

EVALUATION RESULTS

Appropriateness of Flagged Items

REVISE analyzed 572 sentences in the 36 documents and flagged 590 separate items. The user considered 92.5 percent of the flagged items appropriate in the context of the sentences in which they occurred. Of the total, 79.2 percent were single words and 13.3 percent were phrases. The flagged phrases included some passive constructions. Overall, only 7.5 percent of the flagged items were judged inappropriate. On average, REVISE detected about one problem per sentence.

Document Modifications

We used the style program from the WWB (Cherry & Vesterman, 1981) to analyze characteristics of the 36 original and 36 revised documents. Comparisons were made between the original and revised documents on the different measures. Comparisons on some measures were based on very explicit predictions, while others were not. The differences are described below.

Length Characteristics. Table 1 presents means for document, sentence, and word length of the two groups of documents. As expected, the revised documents were shorter overall and, thus, had a shorter average sentence length. (The average number of sentences per document remained constant over the two versions; so, when average document length decreased, so did average sentence length. Both values are tabled for descriptive purposes.) Also as expected, the revised documents had a shorter average word length. All differences were reliable ($p < .001$).

Table 1

Document, Sentence, and Word Length Data for Original and Revised Documents

Item	Mean Number of Items per Unit		
	Original	Revised	Difference
Words/document	340.2	336.7	3.5 *
Words/sentence	22.22	22.00	0.22*
Letters/word	5.34	5.18	0.16*

*$p < .001$.

Readability Measures. Table 2 presents the reading grade level (RGL) computed for each original and revised document by using the Automated Readability Index, Coleman-Liau, Flesch Reading Grade Level, and Kincaid formulas. The revised documents have lower RGLs on all four measures and this was corroborated statistically ($p < .001$). The RGLs of the original and revised documents differ by about 1 year.

Table 2

Reading Grade Level for Original and Revised Documents

Readability Measure	Mean Reading Grade Level (RGL)		
	Original	Revised	Difference
Automated Readability Index	14.81	13.97	0.84*
Coleman-Liau	14.26	13.33	0.93*
Flesch Reading Grade Level	15.82	14.96	0.86*
Kincaid	14.64	13.66	0.98*

*$p < .001$.

Word Type Information. Table 3 presents three groups of specific word measures for the original and revised documents. The first three measures represent three major word classes--nouns, adjectives, and adverbs--as the mean percentage of the total number of words in the document. The next three measures show specific verb types--auxiliary verbs, infinitives and forms of the verb to be--as the mean percentages of the total number of verbs in the document. The final verb category shows passive verbs as a mean percentage of the finite verbs in the document. The next five measures give the mean percentage of the total number of words by other word types--conjunctions, pronouns, nominalizations, prepositions, and abstract words. The final category lists the percentage of sentence beginnings that are expletives.

Table 3

Word Type Information for Original and Revised Documents

Item	Mean Percentage of Total Items in Document		
	Original	Revised	Difference
Word Class			
Noun	31.10	31.04	0.06
Adjective	21.23	21.29	-0.06
Adverb	2.18	2.04	0.14
Verb Types			
Auxiliary[a]	20.86	21.08	-0.22
Infinitive[a]	23.06	23.11	-0.05
To be[a]	42.14	39.69	2.45*
Passive[b]	33.33	31.19	2.14*
Word Types			
Conjunction	3.51	3.64	-0.13*
Pronoun	2.20	2.33	-0.13*
Nominalizations	3.75	3.47	0.28*
Prepositions	12.47	12.49	-0.02
Abstract words	3.14	3.19	-0.05
Expletive beginnings[c]	3.17	0.78	2.39*

[a]Figures are mean percent of all the verbs in the document.

[b]Figures are mean percent of finite verbs in the document.

[c]Figures are mean percentage of all sentence beginnings in the document.

*$p < .005$ (see text).

Specific predictions were made about expletive sentence openers, "to be," and passive verbs. Thus, planned comparisons were justified for these three measures. Specific predictions were not made for the other 10 measures in the table; thus, we observed

caution with respect to the statistical comparisons. Since many comparisons were made, we selected a conservative significance level, $p < .005$, that ensured that the comparison-wide error rate (over 10 comparisons) was $p < .05$ (cf. Keppel, 1973). This significance level was used for all comparisons for which we did not make a specific prediction.

The results from the first three measures show relatively small differences between the documents. None of these differences was reliable.

The verb type results show small differences in auxiliary and infinitive verb use; neither difference approached significance. As expected, the revised documents showed decreased use of forms of the verb "to be" and passive verbs; both differences were reliable in planned comparisons ($p < .005$).

For the final six measures in Table 3, the results show that the revised documents had an increased use of conjunctions and pronouns. While the differences were relatively small in magnitude, they were statistically reliable. The revised documents also contained fewer nominalizations and this difference was reliable. The original and revised documents contained virtually the same percentage of prepositions and abstract words. Finally, as expected, the revised documents had a lower proportion of expletive sentence beginnings than the original documents and this difference was reliable in a planned comparison ($p < .005$).

Sentence Type Information. Table 4 shows sentence-type information for both groups of documents. Sentence types include: simple (one verb and no dependent clause), complex (one independent and one dependent clause, each with one verb), compound (more than one verb and no dependent clause), compound-complex (several dependent clauses or one dependent clause and a compound verb in either the dependent or independent clause). The values presented are mean percentages of the total number of sentences in the document. Since specific predictions were not made about the measures in Table 4, a conservative ($p < .013$) significance level was chosen that kept the comparison-wide error rate to $p < .05$ (over four comparisons). All differences between document types were small and none approached significance.

Table 4

Sentence Type Information for Original and Revised Documents

Sentence Type	Mean Percentage of Sentences in Document		
	Original	Revised	Difference
Simple	57.39	57.14	0.25
Compound	8.00	8.06	-0.06
Complex	26.22	26.03	0.19
Compound-complex	8.58	9.00	0.42

DISCUSSION

Improvements in Writing Quality

The results from the document modification procedure suggest that REVISE altered characteristics of the original documents that lead to improvements in writing and communication efficiency. The REVISE program was not designed specifically to influence the entire range of measures studied; however, the fact that some measures were influenced and others were not is important. Some of the implications of these findings are discussed below.

First, as expected, because of the reduction in average sentence and word length, the RGLs of the revised documents decreased by about 1 year. This decrease, by itself, is not too significant in view of the well-known difficulties with the interpretation (and significance of) readability formulas (Duffy, 1985; Kintsch & Miller, 1984). Nonetheless, the plain English orientation of the REVISE program affected the readability measures in the desired direction.

The specific word type usage results uncovered a variety of interesting patterns. In general, REVISE influenced only the measures considered important to writing quality and communication efficiency and not those that are not. For example, REVISE did not significantly influence the proportion of nouns, adjectives, or adverbs. Simply increasing the proportion of adjectives in a document would not necessarily improve the writing.

With regard to verb forms, as expected the revised documents contained fewer forms of "to be" and passive verbs, both of which are related to the use of passive sentence constructions. Nearly all writing guidelines recommend using active rather than passive constructions.

The revised documents contained fewer nominalizations and expletive beginnings, as writing guidelines recommend. Expletive sentence beginnings typically emphasize the object of the sentence rather than its subject. The revised documents contained more conjunctions, which can add to parallelism in a text, and pronouns, which provide cohesiveness and back reference. Writing guidelines consider that both of these factors improve writing.

REVISE did not affect the use of abstract words. Typically, writing guidelines encourage using concrete as opposed to abstract words.

Rothkopf (1976) pointed out that one document cannot be judged superior to another document without regard to some external referent such as the purpose for which the reader is to use the document. In other words, to determine document improvement, we must consider how the reader will use the document in addition to comparing the surface characteristics of two documents.

Dimensions of Efficient Communication

A critical question raised in research of this sort is whether or not the changes made actually influence the comprehensibility of the document. In fact, an implicit assumption made by authors of writing guidelines is that the guidelines do influence text comprehensibility. The primary guidelines that formed the basis of REVISE involved the use of simpler English and to a lesser extent active sentence constructions. The simpler English principle translates into using words that are generally more familiar to the reader and are thus more frequently used in the English language. REVISE recommends high

frequency words over low frequency ones. A host of findings exist in the psychological literature that show that word familiarity is a powerful determinant of word recognition (Carr, Posner, Pollatsek, & Snyder, 1979). Familiar words are recognized faster and more accurately than unfamiliar words. Similarly, the psycholinguistic literature supports the active writing principle and shows that passive constructions are more difficult for people to understand than are their active counterparts (Coleman, 1964, 1965; Tannenbaum & Williams, 1968), although exceptions exist (Perfetti & Goldman, 1975).

While we can build a case for the benefits of simpler words and phrases and active sentence constructions from the psychological literature, the advantages are found at a fairly low level in the reading process. That is, the beneficial effects are found at the word or sentence level, but are harder to find at higher levels. Duffy and Kabance (1982) have conducted some research with Navy recruits bearing on this issue. The research attempted to improve text by simplifying vocabulary and reducing sentence length. The researchers used these methods to alter several passages from a standard reading test. Among other things, they reduced sentence length from an average of 21.3 to 9.8 words. The revisions failed to produce any meaningful improvement in comprehension on a variety of reading-to-do tasks. Four experiments produced meaningful improvement only for low ability readers on a reading-to-learn task and then only when the vocabulary was simplified. None of the experiments showed an effect for reduced sentence length.

In a similar but more extensive study, Duffy, Curran, and Sass (1983) were interested in comprehension performance on nine pages from a technical manual describing a Navy radio frequency amplifier. The passage had a functional block diagram, a variety of schematics, and roughly 3,000 words of narrative text. Three independent contractors, selected for their expertise in document design, were told to modify the document in any way to maximize comprehension. The contractors made a variety of significant changes. The revisions looked quite different from the original and each other. One contractor spent over 400 man-hours on the revision. Another increased the passage length to over 30 pages. The revisions involved breaking the material up into smaller units, requiring the use of many headings, increased white space, adding new tables and figures, and simplifying the schematics and the block diagrams.

The researchers used two different types of comprehension tests to evaluate the different redesigns. One was a short answer fact test in which subjects could refer back to the text to look up information. The other was a multiple choice test that required the subjects to infer information from what they read. The subjects were electronic technician trainees who had been through roughly six months of electronic technical training and therefore were not novices. None of the revisions facilitated either accuracy of performance on the two comprehension tests or speed of performance in completing the tests.

The primary point in this discussion is that human readers are robust; that is, readers are able to apprehend the deeper meaning of a text in the face of significant departure from "optimal writing." Thus, the low-level changes made by REVISE are unlikely to have a significant effect on the comprehensibility of a text. Indeed in the studies just reviewed, the researchers made major alterations to texts with little or no demonstrable effects on comprehensibility.

Despite the robustness of human readers, some good reasons exist pertaining to why the changes made by REVISE are important. Consider a document that contains many misspelled words. The misspellings alone would not cause the reader to misunderstand the document. Yet, the reader might question the professionalism, thoroughness, or diligence of the author who wrote the document and did not check the spelling. Similarly, the

reader of a document who encounters a number of unusual or awkward words and phrases might question the author's command of the English language or consider that the author is being overly complex and pompous. Factors like these could influence the reader's perception of the document.

Perhaps the most important reason to implement the changes made by REVISE stems from the external pressure to do so. Government writing has been criticized for many years and should be improved where possible. Indeed, the Secretary of the Navy has strongly advocated this practice. The REVISE system is now in use by a technical writing department under contract to the Navy on the East Coast. The system is being used in connection with a project that involves the revision and rewrite of an extremely long (roughly 10,000 pages) technical manual. REVISE is used to assist the technical writers and the editorial staff. Importantly, the system is being used to supplement human capability. In the past, people were hired to perform some of REVISE's tasks. This scenario will become more prevalent as we develop more sophisticated aiding technology.

CONCLUSIONS

REVISE improved the communication efficiency of 36 Navy documents. The system described here represents an efficient computerized writing assistant that can help a writer follow plain English guidelines.

FUTURE COMPUTER-BASED WRITING ASSISTANTS

Future computer-based writing aid systems should be based on an explicit model of comprehension derived from empirical findings from the psycholinguistic literature and have a good, efficient user interface. The psycholinguistic literature tells us a great deal about how people process words, sentences, and paragraphs. From these results, we can derive an explicit model of human text comprehension and incorporate this model into a computerized writing assistant. Such a system would presumably have difficulty processing a section of text just as a person would and could inform the writer of the difficulty. Thus, such a system would not have to be a full-blown natural language comprehension system, but rather a system that would have difficulty where a human reader would.

Kieras (1985) is developing a system like this at the University of Michigan. His system borrows components from earlier simulation work (Kieras, 1982; 1983), is based on a psychological theory of text comprehension (Kintsch & van Dijk, 1978), and includes an augmented transition network (ATN) parser (Woods, 1970), a production system interpreter, and a semantic data base modeled after the Anderson (1976) ACT model. Among other things, the system can determine whether a noun phrase in a sentence is new information (information that has not been mentioned in previous text) or given information (information that has been mentioned earlier) (Clark & Haviland, 1977; Haviland & Clark, 1974). For example, if the author of a technical text refers to a component as the main shipboard power supply and later refers to it as the main supply, the Kieras system would detect the inconsistent terminology, just as a person would.

In order to be maximally useful to a writer, the output of a writing assistant should be easy and efficient to use. The output produced by REVISE has these user friendly characteristics. Thus, an appropriately designed future writing system should combine the model-based, comprehensibility notions of Kieras (1985) with the interface characteristics found in the REVISE program.

REFERENCES

Anderson, J. R. (1976). Language, memory and thought. Hillsdale, NJ: Lawrence Erlbaum Associates.

Carr, T. H., Posner, M. I., Pollatsek, A., & Synder, C. R. R. (1970). Orthography and familiarity effects in word processing. Journal of Experimental Psychology: General, 108, 389-414.

Cherry. L. L. (1978). PARTS-A system for assigning word classes to English text (Computing Science Technical Report No. 81). Murray Hill, NJ: Bell Laboratories.

Cherry, L. L. (1982). Writing tools. IEEE Transactions on communications, 30, 100-105.

Cherry, L. L., & Vesterman, W. (1981). Writing tools--The STYLE and DICTION programs (Computing Science Technical Report No. 91). Murray Hill, NJ: Bell Laboratories.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given/new contact. In R. O. Freedle (Ed.), Discourse processes: Advances in research and theory (pp. 1-40). Norwood, NJ: Ablex Publishing Corporation.

Coleman, E. B. (1964). The comprehensibility of several grammatical transformations. Journal of Applied Psychology, 48, 186-190.

Coleman, E. B. (1965). Learning of prose written in four grammatical transformations. Journal of Applied Psychology, 49, 332-341.

Department of the Air Force. (1973). Guide for Air Force writing (AF Pamphlet 13-2). Washington, DC: Author.

Department of the Interior. (1978). Gobbledygook has gotta go. Washington, DC: Author, Bureau of Land Management.

Department of the Treasury. (1969). Effective revenue writing (Training 9960-12 (8-69)). Washington, DC: Author, Internal Revenue Service.

Department of the Treasury. (1975). Effective writing (Training 9931-01 (Rev. 5-75)). Washington, DC: Author, Internal Revenue Service.

Duffy, T. M. (1985). Readability formulas: What is the use? In T. Duffy & R. Waller (Eds.), Designing usable texts (pp. 113-143). New York: Academic Press.

Duffy, T. M., & Kabance, P. (1982). Testing a readable writing approach to text revision. Journal of Educational Psychology, 74, 733-748.

Duffy, T. M., Curran, T. E., & Sass, D. (1983). Document design for technical job tasks: An evaluation. Human Factors, 25, 143-160.

Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior, 13, 512-521.

Hayes, J. R. (August 1984). Personal communication. San Diego, California.

Heidorn, G. E., Jensen, K., Miller, L. A., Byrd, R. J., & Chodorow, M. S. (1982). The EPISTLE text-critiquing system. IBM System Journal, 21, 305-326.

Keppel, G. (1973). Design and analysis: A researcher's handbook. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Kieras, D. E. (1982). A model of reader strategy for abstracting main ideas from simple technical prose. Text, 2, 47-82.

Kieras, D. E. (1983). A simulation model for the comprehension of technical prose. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 17, pp. 39-80). New York: Academic Press.

Kieras, D. (1985). The potential for advanced computerized aids for comprehensible writing of technical documents (Technical Report No. 17). Ann Arbor, MI: University of Michigan.

Kincaid, J. P., Aagard, J. A., & O'Hara, J. W. (March 1980). Development and test of a computer readability editing system (CRES) (TAEG Report No. 83). Orlando, FL: Training Analysis and Evaluation Group.

Kincaid, J. P., Aagard, J. A., O'Hara, J. W., & Cottrell, L. K. (1981). Computer readability editing system. IEEE Transactions of Professional Communications, 24, 38-41.

Kincaid, J. P., Cottrell, L. K., Aagard, J. A., & Risley, P. (March 1981). Implementing the computer readability editing system (CRES) (TAEG Report No. 98). Orlando, FL: Training Analysis and Evaluation Group.

Kintsch, W., & Miller, J. R. (1984). Readability: A view from cognitive psychology. (1984). In J. Flood (Ed.), Understanding reading comprehension (pp. 220-232). Newark, DE: International Reading Association.

Kintsch, W., & van Dijk, T. A. (1978). Toward a theory of discourse comprehension and production. Psychological Review, 85, 363-394.

Kucera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 293-323.

Macdonald, N. H. (1983). The UNIX Writer's Workbench software: Rationale and design. The Bell System Technical Journal, 62, 1891-1908.

Macdonald, N. H., Frase, L. T., Gingrich, P. S., & Keenan, S. A. (1982). The Writer's Workbench: Computer aids for text analysis. IEEE transactions on Communications, 30, 105-110.

Norman, D. A. (1981). The trouble with UNIX. Datamation, 27, 139-150.

Perfetti, C. A., & Goldman, S. R. (1974). Thematization and sentence retrieval. Journal of Verbal Learning and Verbal Behavior, 6, 335-354.

Perfetti, C. A., & Goldman, S. R. (1975). Discourse functions of thematization and topicalization. Journal of Psycholinguistic Research, 4, 257-271.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), Information processing and cognition: The Loyola symposium (pp. 55-85). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rothkopf, E. Z. (1976). Writing to teach and reading to learn: A perspective on the psychology of written instruction. In Seventy-fifth Yearbook of the National Society for the Study of Education (Part 1, pp. 91-129).

Secretary of the Navy. (August 1983). Navy correspondence manual (SECNAVINST 5216.5C). Washington, DC: Author.

Simply Stated, No. 47. (July 1984). Washington, DC: American Institutes for Research, Document Design Center.

Simply Stated, No. 50. (October 1984). Washington, DC: American Institutes for Research, Document Design Center.

Tannenbaum, P. H., & Williams, F. (1968). Generation of active and passive sentences as a function of subject or object nouns. Journal of Verbal Learning and Verbal Behavior, 7, 246-250.

Woods, W. A. (1970). Transition network grammars for natural language analysis. Communications of the ACM, 13, 591-606.

Wright, P. (1985). Editing policy and processes. In T. Duffy & R. Waller (Eds.), Designing usable texts (pp. 63-96). New York: Academic Press.

DISTRIBUTION LIST

Secretary of the Navy
Military Assistant for Training and Personnel Technology (OSUDS) (R&AT)
Chief of Naval Operations (OP-01B7) (2), (OP-11), (OP-110)
Director of Navy Laboratories (ONT 0722)
Commanding Officer, Service School Command, San Diego (Code 3200)
Chief of Naval Research (Code 442PT)
Chief of Naval Education and Training (Code 00), (Code 00A), (Code N-21)
Director, Naval Education and Training Program Development Center, Great Lakes
Naval Education and Training Support Center, Pacific
Chief of Naval Technical Training (Code 00)
Commander, David W. Taylor Naval Ship Research and Development Center
Commanding Officer, Naval Training Equipment Center (Technical Library) (5), (Code 10)
Commander, Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA (PERI-POT-I), (PERI-SZ)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (AFHRL/MO), (Manpower and Personnel Division), (TSRL/Technical Library FL 2870)
Commander, OPSTNGDIV Air Force Human Resources Laboratory, Wright-Patterson Air Force Base (AFHRL/OT)
Commander, Air Force Human Resources Laboratory, Williams Air Force Base (AFHRL/LR-TDC)
Commander, Air Force Human Resources Laboratory, Lowry Air Force Base (AFHRL/ID)
Superintendent, Naval Postgraduate School
Director of Research, U.S. Naval Academy
Center for Naval Analyses
Defense Technical Information Center (DDAC) (2)

END

FILMED

5-86

DTIC